



TITLE:

目標のエネルギー補充戦略のある 探索ゲーム(モデリングと最適化の 理論)

AUTHOR(S):

宝崎, 隆祐; 池田, 圭子

CITATION:

宝崎, 隆祐 ...[et al]. 目標のエネルギー補充戦略のある探索ゲーム(モデリングと最適化の理論). 数理解析研究所講究録 2006, 1526: 223-231

ISSUE DATE:

2006-12

URL:

<http://hdl.handle.net/2433/58857>

RIGHT:

目標のエネルギー補充戦略のある搜索ゲーム

防衛大学校・情報工学科 宝崎隆祐 (Ryusuke Hohzaki)

Department of Computer Science, National Defense Academy

防衛庁・防衛局 池田圭子 (Keiko Ikeda)

Bureau of Defense Policy, Defense Agency

1 はじめに

通常、搜索ゲームには搜索者と目標のふたりのプレイヤーが参加し、搜索者は目標を見つけるように、目標は搜索者から見つからないようにゲームをプレイする。これまでの研究には、搜索者の戦略が異なる2つのタイプの搜索ゲームモデルが考えられている。いずれにおいても目標の戦略は搜索空間上で移動することであるが、搜索・逃避ゲームと呼ばれる搜索ゲーム [1, 5, 2] においては、搜索者は目標と同じく搜索空間上を移動しながら目標の探知・発見に努めるが、搜索割当ゲームと呼ばれるゲーム [4, 3] における搜索者の戦略は、手持ち搜索資源の搜索空間内への投入である。一方近年の搜索割当ゲームの研究において、目標側のエネルギー制約を加味し現実的な移動環境を考慮した研究 [6, 4, 3] がなされているが、そこで得られた知見からは、エネルギー制約が目標側の機動性に大きく左右し、その結果ゲームの帰趨に少なからず影響を与えることが明らかになった。車、船その他の移動ビークルを目標と考えるとき、エネルギーが移動の制約となるのみならず、エネルギーを移動途中で補充することは普通に行われる行為である。したがって、エネルギー補充を積極的に目標戦略として取り込むことにより、ゲームの結果に影響を与えようとする意図を、目標側は常に持っていると言ってよい。

以上のことから、この研究では、エネルギー補充を目標戦略に加味した場合の搜索割当ゲームをモデル化し、その解法を提案することにより、エネルギー補充に関する最適戦略の特徴について分析することを目的とする。

2 エネルギー補充とモデル

ここではエネルギー補充戦略が組み込まれた搜索割当ゲームのモデルを説明する。

- (A1) 搜索空間は離散地理空間と離散時間空間から成る。地理空間は K 個のセル空間 $K = \{1, \dots, K\}$ で表され、時間空間は T 時点の集合 $T = \{1, \dots, T\}$ で表現される。
- (A2) 目標は搜索空間上を時間とともに移動するが、その移動には次の制約がある。まず、初期時点 $t=1$ での出発セルはセル群 $S_0 \subseteq K$ から選ばれる。時点 t 、セル i から次時点 $t+1$ で移動可能なセル群は $N(i, t)$ である。また、セル i からセル j への移動にはエネルギー $\mu(i, j)$ を消費する。初期時点で保有するエネルギーを e_0 とし、残存エネルギーが無くなれば目標は現在地点から移動できない。
目標はエネルギーを補充することもできる。ある時点においてエネルギーを補充する戦略を採用することにより、次時点当初には目標のエネルギーは δ だけ増加する。ただし、エネルギーの補充戦略が採用されれば、搜索者による目標の被探知確率が一時的に増加する結果となる。その詳細は前提 (A4) で後述する。
目標の1つの移動戦略を $\omega = \{\omega(t), t \in T\}$ で表すが、 $\omega(t) \in K$ は時点 $t \in T$ における目標の存在セルを示す。また、そのエネルギー補充戦略を $\sigma = \{\sigma(t), t \in T\}$ で表

す。 $\sigma(t)$ は 0-1 変数であり、値 1 は時点 t でエネルギーを補充することを示すのに対し、0 は通常の状態（エネルギーを補充しない状態）を表す。

- (A3) 目標を感知するため、時点 τ 以降探索者は探索資源を探索空間内に投入する。この探索可能な時間区間を $\hat{T} = \{\tau, \tau+1, \dots, T\}$ で表す。時点 t において使用可能な資源量は、連続的に分割可能な $\Phi(t)$ である。

探索者の探索資源の 1 つの配分戦略を $\varphi = \{\varphi(i, t), i \in K, t \in T\}$ で表すが、 $\varphi(i, t)$ は時点 t でセル i に投入する資源量を示す。

- (A4) 目標戦略 $\rho = (\omega, \sigma)$ と探索者戦略 φ により、探索者は目標を確率 $1 - \exp(-g(\varphi, \rho))$ で感知できるとする。式 $g(\varphi, \rho)$ は目標経路 ω 上に投入できた効果のある探索資源の重み付き総量で表される。この重みは目標のエネルギー補充状態に依存し、 $\sigma(t) = 0$ の場合には目標セル i においては α_i であるのに対し、エネルギー補充をする ($\sigma(t) = 1$) 場合には β_i とする。通常、 $\alpha_i < \beta_i$ である。

目標を感知すると探索者は利得 1 を得、目標は同量を失う。

前提 (A4) から問題は目標感知確率を支払関数とする 2 人ゼロ和ゲームである。以下においてその定式化を議論してゆく。

エネルギー補充戦略の値域は $S \equiv \{0, 1\}$ である。また、目標戦略ではないが、目標状態を表すため、時点 t の当初における残存エネルギーを表す変数 $e(t)$ を用いる。前提 (A2) にあるように目標戦略は、経路戦略 ω とエネルギー補充戦略 σ の組 $\rho = (\omega, \sigma)$ で与えられるが、それをパスと呼ぶ。ここでパスの実行可能性条件を式により表現すると次のようになる。

- 初期セル条件: $\omega(0) \in S_0$
- 地理的移動制約: $\omega(t+1) \in N(\omega(t), t), t = 2, \dots, T$
- 初期エネルギー制約: $e(1) = e_0$
- エネルギー保存式: $e(t+1) = e(t) - \mu(\omega(t), \omega(t+1)) + \delta\sigma(t), t = 1, \dots, T-1$
- 移動エネルギー制約: $\mu(\omega(t), \omega(t+1)) \leq e(t), t = 1, \dots, T-1$

これらをすべて満足するパスの集合が目標の実行可能パス集合であり、これを P で表わす。前提 (A3) から、探索者の戦略 φ の実行可能領域 Ψ は次式で与えられる。

$$\Psi = \{\varphi \mid \sum_{i \in K} \varphi(i, t) \leq \Phi(t), \varphi(i, t) \geq 0, i \in K, t \in \hat{T}\}. \quad (1)$$

ここで探索者の戦略 φ と目標戦略 $\rho = (\omega, \sigma) \in P$ に対する支払関数を与えよう。時点 t において目標はセル $\omega(t)$ におり、そこに投入した探索資源量 $\varphi(\omega(t), t)$ が感知に有効な資源であるが、その重みは探索者がエネルギー補充戦略 $\sigma(t) = 1$ をとれば $\beta_{\omega(t)}$ であり、未補充戦略をとれば $\alpha_{\omega(t)}$ である。したがって、探索時間全体における有効資源量は次式となる。

$$g(\varphi, \rho) = \sum_{t \in \hat{T}} \varphi(\omega(t), t) \{\alpha_{\omega(t)}(1 - \sigma(t)) + \beta_{\omega(t)}\sigma(t)\}. \quad (2)$$

このとき、前提 (A4) から支払関数である感知確率は次式で表される。

$$R(\varphi, \rho) = 1 - \exp \left(- \sum_{t \in \hat{T}} \varphi(\omega(t), t) \{\alpha_{\omega(t)}(1 - \sigma(t)) + \beta_{\omega(t)}\sigma(t)\} \right). \quad (3)$$

いま、目標に関しては、パス ρ を選択する確率を $\pi(\rho)$ とする混合戦略 $\pi = \{\pi(\rho), \rho \in P\}$ をとることとし、探索者に関しては純粋戦略 φ をとった場合の期待支払は $R(\varphi, \pi) = \sum_{\rho} \pi(\rho) R(\varphi, \rho)$ となる。(3) 式から、この期待支払は変数 φ に関しては凹であり、 π に関しては線形であるから、このような関数のマックスミニ値とミニマックス値が一致することはすでに知られており、その値がゲームの値を与える。以下では、ゲームの値及びプレイヤーの最適戦略を導出する方法を求める。

3 均衡解

ここでは均衡解を求める2つの解法を提案するが、第2の解法は、目標のパス総数が膨大になった場合に対処できるものである。

3.1 パス戦略による均衡解の導出

目標の混合戦略 π に関し $\sum_{\rho \in P} \pi(\rho) = 1$ が成立することに注意すると、期待利得のマックスミニ問題は次のように変形できる。

$$\begin{aligned} \max_{\varphi \in \Psi} \min_{\pi \in \Pi} R(\varphi, \pi) &= \max_{\varphi \in \Psi} \min_{\pi \in \Pi} \sum_{\rho} \pi(\rho) R(\varphi, \rho) = \max_{\varphi \in \Psi} \min_{\rho \in P} R(\varphi, \rho) \\ &= \max_{\varphi \in \Psi, \zeta} \{ \zeta \mid 1 - \exp(-g(\varphi, \rho)) \geq \zeta, \rho \in P \}. \end{aligned}$$

ここで、 $\eta = \ln(1/(1 - \zeta))$ の置き換えによる更なる変形が次のように可能である。

$$\begin{aligned} &= \max_{\varphi \in \Psi, \eta} \{ 1 - \exp(-\eta) \mid g(\varphi, \rho) \geq \eta, \rho \in P \} \\ &= 1 - \exp(-\max_{\varphi \in \Psi, \eta} \{ \eta \mid g(\varphi, \rho) \geq \eta, \rho \in P \}) \\ &= 1 - \exp \left(-\max_{\varphi \in \Psi, \eta} \left\{ \eta \mid \sum_{t \in \hat{T}} \varphi(\omega(t), t) \{ \alpha_{\omega(t)}(1 - \sigma(t)) + \beta_{\omega(t)} \sigma(t) \} \geq \eta, (\omega, \sigma) \in P \right\} \right). \end{aligned} \quad (4)$$

結局、マックスミニ問題は次の線形計画問題を解くことと同値であることが分かる。

$$P^S : \max_{\varphi, \eta} \eta \quad (5)$$

$$s.t. \quad \sum_{t \in \hat{T}} \varphi(\omega(t), t) \{ \alpha_{\omega(t)}(1 - \sigma(t)) + \beta_{\omega(t)} \sigma(t) \} \geq \eta, (\omega, \sigma) \in P \quad (5)$$

$$\sum_{i \in K} \varphi(i, t) \leq \Phi(t), t \in \hat{T} \quad (6)$$

$$\varphi(i, t) \geq 0, i \in K, t \in \hat{T}. \quad (7)$$

上記問題 (P^S) の最適値 η^* を用いれば、原問題のゲームの値は $1 - \exp(-\eta^*)$ により計算できる。同時に、問題 (P^S) は探索者の最適戦略 φ^* を与える。

ところで、(4) 式から、上記問題は線形な期待支払 $R(\varphi, \pi) = \sum_{\rho} \pi(\rho) g(\varphi, \rho)$ をもつマックスミニ問題に他ならないから、以後この線形式を期待支払として取り扱う。次にミニマッ

クス問題を解くため、まず期待支払 $R(\varphi, \pi)$ に次のように変形しよう。

$$\begin{aligned}
 R(\varphi, \pi) &= \sum_{\rho=(\omega, \sigma) \in P} \pi(\rho) \sum_{t \in \hat{T}} \varphi(\omega(t), t) \{ \alpha_{\omega(t)}(1 - \sigma(t)) + \beta_{\omega(t)}\sigma(t) \} \\
 &= \sum_{t \in \hat{T}} \sum_{i \in K} \sum_{\rho=(\omega, \sigma) \in P} \delta_{i\omega(t)} \pi(\rho) \varphi(i, t) \{ \alpha_i(1 - \sigma(t)) + \beta_i\sigma(t) \} \\
 &= \sum_{t \in \hat{T}} \sum_{i \in K} \sum_{\rho=(\omega, \sigma) \in P_{it}} \pi(\rho) \varphi(i, t) \{ \alpha_i(1 - \sigma(t)) + \beta_i\sigma(t) \} \quad (8)
 \end{aligned}$$

上式中の記号 δ_{ij} はクロネッカーのデルタを、 P_{it} は時刻 t にセル i を通るパス全体の集合、すなわち $P_{it} \equiv \{ \rho = (\omega, \sigma) \in P \mid \omega(t) = i \}$ を意味する。 $\sum_{i \in K} \varphi(i, t) \leq \Phi(t)$ であることに注意すれば、期待利得の φ に関する最大化は次式を導く。

$$\max_{\varphi \in \Psi} R(\varphi, \pi) = \sum_{t \in \hat{T}} \Phi(t) \max_{i \in K} \sum_{\rho=(\omega, \sigma) \in P_{it}} \pi(\rho) \{ \alpha_i(1 - \sigma(t)) + \beta_i\sigma(t) \}. \quad (9)$$

上式中で $\max_{i \in K}$ 以下の式を $\nu(t)$ で置き換えてやると、ミニマックス問題は次のように定式化できる。

$$\begin{aligned}
 P^T : \quad & \min_{\pi, \nu} \sum_{t \in \hat{T}} \Phi(t) \nu(t) \\
 \text{s.t.} \quad & \sum_{\rho=(\omega, \sigma) \in P_{it}} \pi(\rho) \{ \alpha_i(1 - \sigma(t)) + \beta_i\sigma(t) \} \leq \nu(t), \quad t \in \hat{T}, i \in K \quad (10)
 \end{aligned}$$

$$\sum_{\rho \in P} \pi(\rho) = 1 \quad (11)$$

$$\pi(\rho) \geq 0, \quad \rho \in P. \quad (12)$$

因みに、(10) 式は $\nu(t)$ への置き換えを正当化するために必要な不等式である。上記問題 (P^T) を解くことにより、目標の最適混合戦略 π^* が得られる。

問題 (P^S) と (P^T) とは同じ最適解を与えることは、両者が双対の関係であることを証明することにより確認できるし、そのような双対性は2人のプレイヤーの最適戦略を求める際にも1つの問題だけを解けばよいことを我々に教える。以上の結果は次の定理でまとめられる。

定理 1 ゲームの値は線形計画問題 (P^S)、あるいは (P^T) を解くことにより得られる。また、探索者の最適戦略 φ^* は、問題 (P^S) の最適解、あるいは問題 (P^T) の制約条件式 (10) に対応する双対変数の最適値により与えられる。また、目標の最適混合戦略 π^* は、問題 (P^T) の最適解、または問題 (P^S) の条件式 (5) に対応する最適な双対変数により与えられる。

3.2 マルコフ移動戦略による均衡解の導出

条件式 (5), (12) から分かるように、定式化 (P^S) 及び (P^T) では、目標のすべての実行可能なパスを数え挙げている。ところが、パスの実行可能性に何ら制約がない場合を考えると、その総数は $(2|K|)^{|\hat{T}|}$ となり、探索空間のサイズが大きな場合には、いわゆる組合せ的爆発が問題の現実的な解法には障害となる。ここでは目標戦略をパス戦略からマルコフ移動戦略に変更することにより、そのような障害を回避する工夫を考えよう。

話を簡単にするため、エネルギー消費関数 $\mu(i, j)$ 、初期エネルギー量 e_0 及び補充エネルギー量 δ は整数値をとるとする。すると、可能なエネルギー状態は $E = \{0, \dots, e_0, \dots, e_0 + T\delta\}$ で表すことができる。目標の状態は、いつ (t) 、どのセルに (i) 、どれだけ残存エネルギーをもって (e) 、どのようなエネルギー補充状態であるか (s) を示す4つの値の組 (i, e, s, t) で表現できる。ただし、エネルギー補充状態は $s = \{0, 1\}$ である。パス選択の混合戦略 $\pi(\rho)$ 、 $\rho \in P$ 、は状態 (i, e, s, t) にある目標の存在確率を形成するから、この確率を $q(i, e, s, t)$ とし、この状態から次の時点 $t+1$ でセル j 、エネルギー補充状態 r に移行するマルコフ型の遷移確率を $v(i, e, s, t, j, r)$ として、この2つの変数 q, v を目標の新たな戦略と考えよう。ただし、探索者の戦略は前節と同じ探索資源配分計画 φ で表す。ここで、状態 (i, e, s, t) から次時点で移動可能なセル群は $C(i, e, t) = \{j \in N(i, t) | \mu(i, j) \leq e\}$ であり、状態 (i, e, s, t) に移行可能な前の時点 $t-1$ における存在セルは、そこでのエネルギー補充状態 r に依存した形で $C^*(r, i, e, t) = \{j \in K | i \in C(j, e + \mu(j, i) - \delta r, t-1)\}$ と書くことができる。

ここで、期待利得 (8) 式を更に変形すると次のようになる。

$$\begin{aligned} R(\varphi, \pi) &= \sum_{t \in \hat{T}} \sum_{i \in K} \varphi(i, t) \sum_{\rho=(\omega, \sigma) \in P_{it}} \pi(\rho) \{ \alpha_i (1 - \sigma(t)) + \beta_i \sigma(t) \} \\ &= \sum_{t \in \hat{T}} \sum_{i \in K} \varphi(i, t) \left(\alpha_i \sum_{\rho \in P_{it, \sigma(t)=0}} \pi(\rho) + \beta_i \sum_{\rho \in P_{it, \sigma(t)=1}} \pi(\rho) \right). \end{aligned} \quad (13)$$

最終式における $\sum_{\rho \in P_{it, \sigma(t)=0}} \pi(\rho)$ 及び $\sum_{\rho \in P_{it, \sigma(t)=1}} \pi(\rho)$ は、その意味を考えれば、それぞれ $s = 0$ 及び $s = 1$ とした確率 $\sum_{e \in E} q(i, e, s, t)$ に置き換えることができ、新しい戦略 $q(\cdot)$ を用いれば期待支払は次式となる。

$$R(\varphi, \pi) = \sum_{t \in \hat{T}} \sum_{i \in K} \varphi(i, t) \left(\alpha_i \sum_{e \in E} q(i, e, 0, t) + \beta_i \sum_{e \in E} q(i, e, 1, t) \right). \quad (14)$$

以後、これまでの議論を用いてミニマックス最適化を行ってゆくが、ここでは動的計画法により定式化することにしよう。そのため、時点 t 以降の最適な探索資源配分計画により実現できる期待利得の最大値を $h(t)$ で表す。(14) 式から、時刻 t 、セル i 及びエネルギー未補充の状態 $s = 0$ の場合に生じる期待利得は $\varphi(i, t) \alpha_i \sum_{e \in E} q(i, e, 0, t)$ であり、補充状態 $s = 1$ の場合には $\varphi(i, t) \beta_i \sum_{e \in E} q(i, e, 1, t)$ であるから、 $h(t)$ と $h(t+1)$ の間には、次の漸化式が成り立つ。

$$\begin{aligned} h(t) &= \max_{\varphi \in \Psi} \left\{ \sum_{i \in K} \varphi(i, t) \left(\alpha_i \sum_{e \in E} q(i, e, 0, t) + \beta_i \sum_{e \in E} q(i, e, 1, t) \right) + h(t+1) \right\} \\ &= \Phi(t) \max_{i \in K} \sum_{e \in E} (\alpha_i q(i, e, 0, t) + \beta_i q(i, e, 1, t)) + h(t+1). \end{aligned} \quad (15)$$

探索資源に関する条件 $\sum_{i \in K} \varphi(i, t) \leq \Phi(t)$ により、第1式から (15) 式が導出されるが、この式から不等式 $h(t) \geq \Phi(t) \sum_{e \in E} (\alpha_i q(i, e, 0, t) + \beta_i q(i, e, 1, t)) + h(t+1)$ が任意のセル i に対し成り立つ。また、探索は時点 τ から始まるから探索全体での最大期待利得は $h(\tau)$ であり、期待利得のミニマックス問題は次の問題を解けばよい。

$$\begin{aligned} (P_m^T) \quad &\min h(\tau) \\ \text{s.t.} \quad &h(t) \geq \Phi(t) \sum_{e \in E} (\alpha_i q(i, e, 0, t) + \beta_i q(i, e, 1, t)) + h(t+1), i \in K, t = \tau, \dots, T-1 \end{aligned} \quad (16)$$

$$h(T) \geq \Phi(T) \sum_{e \in E} (\alpha_i q(i, e, 0, T) + \beta_i q(i, e, 1, T)), \quad i \in K \quad (17)$$

$$q(i, e, s, t) = \sum_{r \in S} \sum_{j \in C(i, e, t)} v(i, e, s, t, j, r), \quad i \in K, e \in E, s \in S, t = 1, \dots, T-1 \quad (18)$$

$$q(i, e, s, t) = \sum_{r \in S} \sum_{j \in C^*(r, i, e, t)} v(j, e + \mu(j, i) - \delta r, r, t-1, i, s), \quad i \in K, e \in E, s \in S, t = 2, \dots, T \quad (19)$$

$$\sum_{i \in S_0} \sum_{s \in S} q(i, e_0, s, 1) = 1 \quad (20)$$

$$\sum_{i \in K} \sum_{e \in E} \sum_{s \in S} q(i, e, s, t) = 1, \quad t \in T \quad (21)$$

$$v(i, e, s, t, j, r) \geq 0, \quad i, j \in K, e \in E, s, r \in S, t = 1, \dots, T-1. \quad (22)$$

(18) 式以下の条件は、存在確率 $q(\cdot)$ と遷移確率 $v(\cdot)$ の確率流に関して成り立つ保存則や初期条件を記述したものである。この線形計画問題を解くことにより、目標の最適戦略が存在確率、遷移確率として得られる。

次に探索者の最適戦略を求めるため、ミニマックス問題を線形計画問題として定式化してみよう。 $w(i, e, s, t)$ を、状態 (i, e, s, t) から出発した目標が以後最適パスを選択することにより達成できる最小の支払とする。時点 τ より前では探索は実施されないから、連続時点 $t, t+1 (\leq \tau)$ での $w(\cdot)$ の値の間には次の関係が成り立つ。

$$w(i, e, s, t) = \min_{j \in C(i, e, t)} \min_{r \in S} w(j, e - \mu(i, j) + \delta s, r, t+1). \quad (23)$$

探索が開始される時点 τ 以降においては、時点 t で発生する利得が $(\alpha_i(1-s) + \beta_i s)\varphi(i, t)$ であることに注意し、次の関係式を得る。

$$w(i, e, s, t) = \min_{j \in C(i, e, t)} \min_{r \in S} \{(\alpha_i(1-s) + \beta_i s)\varphi(i, t) + w(j, e - \mu(i, j) + \delta s, r, t+1)\}. \quad (24)$$

目標の初期条件に留意すれば、探索全体における最小利得は $\min_{i \in S_0} \min_{s \in S} w(i, e_0, s, 1)$ であるから、マックスミニ最適化問題は次のように定式化できる。

$$(P_m^S) \quad \max \xi \quad (25)$$

$$\text{s.t. } w(i, e_0, s, 1) \geq \xi, \quad i \in S_0, s \in S$$

$$w(i, e, s, 1) = 0, \quad i \in K, e \in E \setminus \{e_0\}, s \in S$$

$$w(i, e, s, 1) = 0, \quad i \in K \setminus S_0, e \in E, s \in S$$

$$w(i, e, s, t) \leq w(j, e - \mu(i, j) + \delta s, r, t+1),$$

$$i \in K, j \in C(i, e, t), e \in E, s, r \in S, t = 1, \dots, \tau-1$$

$$w(i, e, 0, t) \leq \alpha_i \varphi(i, t) + w(j, e - \mu(i, j), r, t+1),$$

$$i \in K, j \in C(i, e, t), e \in E, r \in S, t = \tau, \dots, T-1$$

$$w(i, e, 1, t) \leq \beta_i \varphi(i, t) + w(j, e - \mu(i, j) + \delta, r, t+1),$$

$$i \in K, j \in C(i, e, t), e \in E, r \in S, t = \tau, \dots, T-1$$

$$w(i, e, 0, T) = \alpha_i \varphi(i, T), \quad i \in K, e \in E$$

$$w(i, e, 1, T) = \beta_i \varphi(i, T), \quad i \in K, e \in E$$

$$\sum_{i \in K} \varphi(i, t) = \Phi(t), \quad t = \tau, \dots, T$$

$$\varphi(i, t) \geq 0, \quad i \in K, t = \tau, \dots, T.$$

パス戦略による2つの問題(P^S), (P^T)と同様, 問題(P_m^S)と(P_m^T)の間にはお互いに対する双対性が存在するが, 紙数の関係上その証明は省略する. ここで, 目標戦略をマルコフ移動とした場合のゲームの解の導出について定理としてまとめておこう.

定理 2 ゲームの値は, 線形計画問題(P_m^T)または(P_m^S)を解くことにより得られる. 探索者の最適戦略 φ^* は, 問題(P_m^S)の最適解, あるいは問題(P_m^T)の条件式(16)及び(17)に対する最適双対変数により与えられ, 目標の最適戦略 $q^*(\cdot)$, $v^*(\cdot)$ は, 問題(P_m^T)の最適解により与えられる.

4 数値例

ここでは, 探索空間 $K \times T = \{1, \dots, 9\} \times \{1, \dots, 7\}$ において時刻 $t = 1$ にセル $S_0 = \{1\}$ から出発する目標と, それを探知しようと時刻 $\tau = 2$ から探索を開始する探索者との探索ゲームを考える. セル空間はセル番号 $1, 2, \dots, 9$ の順に連続して接しており, 目標はエネルギーが許せば3隣接セルまで移動可能である. すなわち, $N(i, t) = \{i-3, i-2, \dots, i, \dots, i+3\} \cap K$ とする. 目標は初期エネルギー $e_0 = 4$ を持つが, セル間の移動には距離の自乗に比例するエネルギー $\mu(i, j) = |i - j|^2$ を消費する. 目標は移動戦略とともに, エネルギー未補充($s = 0$)か補充($s = 1$)の戦略を各時点ですることができるが, $s = 0$ の戦略に対する単位探索資源の探知に対する効率はすべてのセルで同じであると仮定し, $\alpha_i = 1$ と設定する. $s = 1$ に対する β_i もすべてのセルで同じ値1.5であるとする. また, $s = 1$ の戦略により目標が補充できるエネルギー量は $\delta = 1$ である.

さて, (14)式から分かるように, $\alpha_i \sum_e q(i, e, 0, t) + \beta_i \sum_e q(i, e, 1, t)$ は時刻 t のセル i に投入する単位探索資源の支払に対する寄与を表すから, この値の大小は探索者にとっての資源投入場所としての魅力を表す指標と考えて良い. この値をセル i , 時刻 t での探知効力と呼ぼう. 一方同じ式から, $\varphi(i, t)$ の大きさは目標にとっての忌避すべき場所の指標を与え, 基本的にはこの値が大きな所にはできるだけ小さな探知効力を, 小さな所には大きな探知効力を形成するように動くべきだと言える.

以上のようなパラメータ設定の下で, ゲームの解を求めた. 表1には, 目標の最適存在確率 $\sum_e \sum_s q(i, e, s, t)$ が, 横軸に時間 t , 縦軸にセル i をとって表示されている. 表2には, この存在確率を2つに分け, エネルギー補充状態での確率 $\sum_e q(i, e, 1, t)$ を上, 未補充の確率 $\sum_e q(i, e, 0, t)$ を下に記述している. また, 探知効力の値が表3に記載されている. 表4は探索者の最適探索資源配分計画である. エネルギー補充戦略 $s = 1$ が目標の探知され易さを招くこのケースでは, 表2から分かるとおり, 補充の戦略をとる回数は少ないが, 目標は存在領域の境界付近 $(t, i) = (1, 1), (3, 3), (3, 4), (4, 5), (5, 6), (6, 6), (6, 7), (7, 8)$ でのみ補充戦略を採ることにより, 存在領域の拡大化を目指している. これは探索者の探索資源投入を広い場所に拡散させ, 各セルへの投入密度を低く押さえる効果がある. また上記の場所でエネルギー補充を選択する確率を比べると, 目標存在領域があまり拡大していない早い時点においてその確率はより大きく, 早い段階における目標存在領域の局限化状態を早期に解消しようとする意図が見られる. また表1では, 探知され易さの観点から, 補充確率のあるセルでのその確率は小さく押さえられている. 結果として, 探知効力は表3のように $(t, i) = (7, 8)$ を除いたすべての場所で同じ値が実現されており, 目標が探知効力の一様性を指向していることも見て取れる.

さて, ほぼ同じ探知効力値の並ぶ表3にもかかわらず, 表4の探索者の探索資源配分を見ると, 目標存在領域内部のセルでは一様配分が基本的性質であるものの, 投入量が他のセルに比べ少ない場所 $(t, i) = (4, 5), (5, 6), (7, 6), (7, 7)$ と多い場所 $(t, i) = (6, 5), (6, 6), (6, 7)$ がある. これは目標移動が時間に伴う連続移動であることから, 存在領域の境界付近を動き

より遠くに逃避しようとする目標に対しては、時刻 $t=6$ での境界領域の重点搜索により対処しようとする搜索者の意図がある。このため、目標はその存在確率を $(t, i) = (6, 6), (6, 7)$ では減少させることになるが、このことが、次の時刻 $t=7$ でのセル 8 の目標存在確率及び探知効力を 0.081 及び 0.095 に低下させ、搜索者はここへの搜索資源投入を節約して他の内部のセルへの資源投入に回すことが可能となっている。

5 おわりに

本論文では、目標がエネルギーの補充戦略をもつ場合の搜索割当ゲームを 2 人ゼロ和ゲームに定式化し、搜索者と目標の最適戦略を導出する線形計画問題への定式化を提案した。数値例による分析の結果、プレイヤーは極めて巧妙に敵対者に対し適応的に最適戦略をとっている様子がわかり、プレイヤーの最適戦略の基本的な特徴が明らかになった。

表 1. 目標の最適存在確率

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.081
7	0	0	0	0	0	0.115	0.131
6	0	0	0	0	0.118	0.108	0.131
5	0	0	0	0.153	0.176	0.155	0.131
4	0	0	0.194	0.212	0.176	0.155	0.131
3	0	0.333	0.268	0.212	0.176	0.155	0.131
2	0	0.333	0.269	0.212	0.176	0.155	0.131
1	1	0.333	0.269	0.212	0.176	0.155	0.131
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

表 2. 目標の最適エネルギー補充戦略

	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
	0	0	0	0	0	0	0.030
8	0	0	0	0	0	0	0.051
	0	0	0	0	0	0.081	0
7	0	0	0	0	0	0.034	0.131
	0	0	0	0	0.118	0.094	0
6	0	0	0	0	0	0.014	0.131
	0	0	0	0.118	0	0	0
5	0	0	0	0.035	0.176	0.155	0.131
	0	0	0.150	0	0	0	0
4	0	0	0.044	0.212	0.176	0.155	0.131
	0	0	0.003	0	0	0	0
3	0	0.333	0.265	0.212	0.176	0.155	0.131
	0	0	0	0	0	0	0
2	0	0.333	0.269	0.212	0.176	0.155	0.131
	0.75	0	0	0	0	0	0
1	0.25	0.333	0.269	0.212	0.176	0.155	0.131
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

表 3. 探知効力

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0.095
7	0	0	0	0	0	0.155	0.131
6	0	0	0	0	0.176	0.155	0.131
5	0	0	0	0.212	0.176	0.155	0.131
4	0	0	0.269	0.212	0.176	0.155	0.131
3	0	0.333	0.269	0.212	0.176	0.155	0.131
2	0	0.333	0.269	0.212	0.176	0.155	0.131
1	1.375	0.333	0.269	0.212	0.176	0.155	0.131
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

表 4. 搜索者の最適搜索資源投入戦略

9	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0
7	0	0	0	0	0.148	0.074	
6	0	0	0	0.138	0.148	0.148	
5	0	0	0.093	0.172	0.148	0.156	
4	0	0.25	0.227	0.172	0.139	0.156	
3	0.333	0.25	0.227	0.172	0.139	0.156	
2	0.333	0.25	0.227	0.172	0.139	0.156	
1	0.333	0.25	0.227	0.172	0.139	0.156	
cells	t=1	t=2	t=3	t=4	t=5	t=6	t=7

参考文献

- [1] J.M. Danskin, A Helicopter versus Submarine Search Game, *Operations Research*, **16**, pp.509–517, 1968.
- [2] J.N. Eagle and A.R. Washburn, Cumulative Search-Evasion Games, *Naval Research Logistics*, **38**, pp.495–510, 1991.
- [3] R. Hohzaki, Search Allocation Game, *European Journal of Operational Research*, **172**, pp.101–119, 2006.
- [4] R. Hohzaki, and A. Washburn, An Approximation for a Continuous Datum Search Game with Energy Constraint, *Journal of the Operations Research Society of Japan*, **46**, pp.306–318, 2003.
- [5] A.R. Washburn, Search-Evasion Game in a Fixed Region, *Operations Research*, **28**, pp.1290–1298, 1980.
- [6] A.R. Washburn and R. Hohzaki, The Diesel Submarine Flaming Datum Problem, *Military Operations Research*, **4**, pp.19–30, 2001.